The Cavity Method in Random Matrix Theory

dated: October 1, 2022

In these notes we will discuss the cavity method, and how it can be used to find the spectral density of a random matrix. We will use it to work out two examples: the semicircle law and the Marchenko-Pastur distribution for the Gaussian orthogonal and Wishart ensembles respectively. Understanding these concepts will need several ingredients. First we introduce the notion of a *resolvent* which is like a generating function for our random matrix, and contains the information necessary to calculate the spectral density. Then we employ the cavity method to derive a self-consistent equation for the resolvent of an $N \times N$ matrix as $N \to \infty$. By solving this equation we can obtain the information we want about the probability density of eigenvalues of various random matrix ensembles.

The discussion here is similar to that from *A First Course in Random Matrix Theory* by Potters and Bouchaud, which I highly recommend as a resource for those curious about the subject.

1 The Resolvent

Here we introduce a very useful function, known as the resolvent (or Greens function) of an $N \times N$ matrix $\mathbf{H}$. The function is defined as

$$g(z) = \left\langle \frac{1}{N} \operatorname{Tr} \frac{1}{z - \mathbf{H}} \right\rangle. \tag{1}$$

Here $1/\mathbf{M}$ is shorthand for $\mathbf{M}^{-1}$, and $z - \mathbf{H}$ is short for $z\mathbf{I} - \mathbf{H}$ where $\mathbf{I}$ is the identity matrix. An alternative representation of the resolvent can be found by expanding it in an infinite series in H. Then one obtains

$$g(z) = \frac{1}{N} \sum_{k=0}^{\infty} \frac{1}{z^{k+1}} \left\langle \operatorname{Tr}(\mathbf{H}^k) \right\rangle, \tag{2}$$

and so we see that $g(z)$ encodes information about all the moments of $\mathbf{H}$. [It turns out that $g(z)$ is also the Laplace transform of the response function of a single variable in the linear dynamical system $\dot{\mathbf{x}} = \mathbf{H}\mathbf{x}$ — and is sometimes called a Greens function.]

It turns out that in the limit that the matrix $\mathbf{H}$ becomes very large, the resolvent becomes equal to

$$g(z) = \int \frac{\rho(z')}{z - z'} dz' \tag{3}$$

where $\rho(z)$ is the density of eigenvalues of the matrix H. This kind of transform is known as a *Stieltjes transform* of the function $\rho(z)$. What's the intuition for this? Note that $z\mathbf{I} - \mathbf{H}$ obtains a 0 eigenvalue whenever z is an eigenvalue of H, so $\operatorname{Tr} \frac{1}{z\mathbf{I}-\mathbf{H}}$ has a pole whenever z is an eigenvalue of $\mathbf{H}$. Similarly $g(z)$ has a pole whenever z is an eigenvalue of $\mathbf{H}$, since then $\rho(z')$ does not vanish though the denominator in the integrand $z - z'$ vanishes. As the matrix $\mathbf{H}$ becomes large, we have an infinite number of these poles at the values of z that are eigenvalues of $\mathbf{H}$. The idea is that in the $N \to \infty$ limit, this is equivalent to weighing the function $\frac{1}{z-z'}$ by the spectral density $\rho(z')$ and integrating over z'.

If you are not convinced, we can show that our definition of the resolvent is equivalent to the Stieltjes transform of the spectral density as follows:

$$\operatorname{Tr} \frac{1}{z - \mathbf{H}} = \sum_i x_i = \int \sum_i \delta(x - x_i) x dx \text{ where } x_i \text{ are eigenvalues of } \frac{1}{z - \mathbf{H}}. \tag{4}$$

These eigenvalues are given by

$$x_i = \frac{1}{z - z_i} \text{ where } z_i \text{ are eigenvalues of } \mathbf{H}. \tag{5}$$

Now we can also define our spectral density as

$$\rho(z) = \sum_i \delta(z - z_i) \implies \int \frac{\rho(z')}{z - z'} dz' = \int \frac{\sum_i \delta(z' - z_i)}{z - z'} dz'. \tag{6}$$

So it remains to be shown that Equation 4 reproduces the definition of spectral density, or in other words

$$\int \sum_i \delta\left(x - \frac{1}{z - z_i}\right) x dx = \int \frac{\sum_i \delta(x - z_i)}{z - x} dx. \tag{7}$$

However this equality is satisfied by the definition of the δ function: both sides evaluate to $\sum_i \frac{1}{z-z_i}$. Therefore we should be convinced of the equivalence between $g(z)$ and the Stieltjes transform of $\rho(z)$.

The next thing to show is that if we can calculate the resolvent, we can calculate the spectral density of the eigenvalues of $\mathbf{H}$. This *spectral density* $\rho(z)$ can be viewed as the marginal distribution of a single eigenvalue of $\mathbf{H}$, over all possible realizations of $\mathbf{H}$ drawn from some underlying ensemble. $\rho(z)$ is a quantity of great interest in the theory of random matrices. How can we calculate it from its Hilbert transform $g(z)$? Note that if we take $g(z - i\epsilon)$ then this expands to

$$g(z - i\epsilon) = \int \frac{\rho(z')}{z - z' - i\epsilon} dz' = \int \frac{\rho(z')(z - z')}{(z - z')^2 + \epsilon^2} dz' + i\epsilon \int \frac{\rho(z')}{(z - z')^2 + \epsilon^2} dz'. \tag{8}$$

Using the fact that

$$\delta(z) = \frac{1}{\pi} \lim_{\epsilon \to 0^+} \frac{\epsilon}{\epsilon^2 + x^2}, \tag{9}$$

we have that

$$\rho(z) = \frac{1}{\pi} \lim_{\epsilon \to 0^+} \operatorname{Im} g(z - i\epsilon). \tag{10}$$

Therefore if we can calculate $g(z)$, we can get $\rho(z)$.

2 Calculating the Spectral Density

We will first derive a useful formula that relates the inverse of a matrix to the inverse of its blocks. Take a matrix

$$\mathbf{M} = \begin{pmatrix} M_{11} & \mathbf{M}_{12} \\ \mathbf{M}_{21} & \mathbf{M}_{22} \end{pmatrix} \tag{11}$$

where M_{11} is the upper left entry, and we have written it in this form to emphasize that $\mathbf{M}$ can be decomposed into blocks. What is the relationship between $\mathbf{M}^{-1}$ and $\mathbf{M}_{22}^{-1}$? To figure this out, we can write $\mathbf{M}^{-1}$ in block diagonal form as well.

$$\mathbf{M}^{-1} = \begin{pmatrix} Q_{11} & \mathbf{Q}_{12} \\ \mathbf{Q}_{21} & \mathbf{Q}_{22} \end{pmatrix} \tag{12}$$

Then the fact that $\mathbf{M}\mathbf{M}^{-1} = \mathbf{I}$ allows us to deduce that

$$Q_{11}^{-1} = M_{11} - \mathbf{M}_{12}(\mathbf{M}_{22})^{-1}\mathbf{M}_{21}. \tag{13}$$

This is the central formula, also known as the *Schur complement formula*, that will allow us to construct a self-consistent equation for the resolvent.

2.1 The Gaussian Orthogonal Ensemble

As a warm up problem we can use this equation to calculate the resolvent and therefore the spectral density of the Gaussian orthogonal ensemble, which is a rotationally invariant gaussian-entry ensemble which is symmetric (and therefore admits diagonalization by an orthogonal matrix). In particular the joint probability density function of the entries H_{ij} of a Gaussian orthogonal matrix is

$$p(\mathbf{H}) = \frac{1}{Z} e^{-\frac{N}{4} \operatorname{Tr}(\mathbf{H}^2)}. \tag{14}$$

Recalling that $\operatorname{Tr}(\mathbf{H}^2) = \sum_{jk} H_{jk}^2$ for a symmetric matrix $\mathbf{H}$, this distribution means that each element of $\mathbf{H}$ is gaussian distributed with variance $1/N$, except for the diagonal entries which have variance $2/N$.

Now let's define the matrix $\mathbf{M} = z - \mathbf{H}$, and $\mathbf{G}(z) = \mathbf{M}^{-1}$ and $\mathbf{G}_2(z) = (\mathbf{M}_{22})^{-1}$. The cavity method leverages the fact that as N gets large, $\langle \operatorname{Tr} \mathbf{G}_2(z) \rangle$ should approach $g(z)$. Now we can use the Schur formula, which tells us that

$$G(z)_{11}^{-1} = M_{11} - \mathbf{H}_{12}\mathbf{G}_2(z)\mathbf{H}_{21}. \tag{15}$$

Taking the expectation of this expression gives

$$\frac{1}{g(z)} = z - \left\langle \sum_j H_{1j}^2 (\mathbf{M}_{22})_{jj}^{-1} \right\rangle = z - g_2(z) = z - g(z). \tag{16}$$

Note that we have used the fact that $\left\langle G(z)^{-1}_{11} \right\rangle = \frac{1}{g(z)}$ which comes from the self-averaging property of the resolvent, such that each of the diagonal elements of $\mathbf{M}^{-1}$ will be the same with $\mathcal{O}(N^{-1/2})$ fluctuations. Then we can solve this algebraic equation for $g(z)$ to obtain

$$g(z) = \frac{1}{2}\left(z \pm i\sqrt{4 - z^2}\right). \tag{17}$$

Using Equation 10, one can find that $\rho(x) = \frac{1}{2\pi}\sqrt{4 - x^2}$. This is the celebrated semicircle law, and is the spectral distribution for many random matrix ensembles, including all the gaussian ensembles.

2.2 The Wishart Ensemble

Consider a Wishart matrix $\mathbf{W} = \frac{1}{T}\mathbf{H}\mathbf{H}^\intercal$, where $\mathbf{H}$ has dimensions $N \times T$, so that $\mathbf{W}$ is an $N \times N$ matrix. The entries of $\mathbf{H}$ are each i.i.d. gaussian random variables, with no symmetry conditions. Defining $\mathbf{M} = z - \mathbf{W}$, our goal is to calculate $g(z) = \left\langle \operatorname{Tr} \mathbf{M}^{-1} \right\rangle$. We will define the matrix $\mathbf{G}(z) = \mathbf{M}^{-1}$ and $\mathbf{G}_2(z) = (\mathbf{M}_{22})^{-1}$. The Schur formula tells us that

$$G(z)^{-1}_{11} = M_{11} - \mathbf{M}_{12}(\mathbf{M}_{22})^{-1}\mathbf{M}_{21}. \tag{18}$$

We see that the resolvent of the submatrix $\mathbf{W}_{22}$ is $g_2(z) = \frac{1}{N}\left\langle \operatorname{Tr}(\mathbf{M}_{22})^{-1} \right\rangle$. In the limit that the size of $\mathbf{W}$ gets very large, the resolvents $g(z)$ and $g_2(z)$ should be equal. This equality allows us to enforce an additional constraint and solve for $g(z)$. Instead of calculating $g_2(z)$ directly, we see that it comes out as part of an expression when we average the Schur formula.

Another way to say this is: our goal is to express $g_2(z)$ in terms of $g(z)$ using the Schur formula. Then at the end we will set $g(z) = g_2(z)$ and be able to solve for this resolvent.

In order to take the expectation of Equation 18, we first note that, just as in the case of the Gaussian orthogonal ensemble, we obtain $\left\langle G(z)^{-1}_{11} \right\rangle = \frac{1}{g(z)}$. Then we note that $\langle M \rangle_{11} = \langle z - W_{11} \rangle = z - 1$.

Now we tackle the trickier term $\mathbf{M}_{12}(\mathbf{M}_{22})^{-1}\mathbf{M}_{21}$, whose expectation we wish to calculate. This term can be written as

$$\sum_{j,k} M_{1j}(\mathbf{M}_{22})^{-1}_{jk} M_{k1}. \tag{19}$$

We now have to do some careful bookkeeping of indices. Noting that the off diagonal components of $\mathbf{M}$ are the same as those of $\mathbf{W}$, and expanding out the components of $\mathbf{W}$ in terms of $\mathbf{H}$, we obtain

$$\frac{1}{T^2}\sum_{j,k=2}^{N}\sum_{s,t=1}^{T} H_{1s}H_{js}(\mathbf{M}_{22})^{-1}_{jk}H_{kt}H_{1t} \tag{20}$$

Under expectation the factor $H_{1s}H_{1t}$ is equivalent to a factor of δ_{ts}. Therefore this term becomes

$$\frac{1}{T^2}\sum_{j,k=2}^{N}\sum_{t=1}^{T} H_{jt}H_{kt}(\mathbf{M}_{22})^{-1}_{jk} = \frac{1}{T}\operatorname{Tr}\mathbf{W}_{22}\mathbf{G}_2(z). \tag{21}$$

We can then simplify this using the definition of $\mathbf{G}_2(z)$ in terms of $\mathbf{W}_{22}$, namely $\mathbf{G}_2(z) = (z - \mathbf{W}_{22})^{-1}$. Therefore

$$-\mathbf{W}_{22}\mathbf{G}_2(z) = (z - \mathbf{W}_{22})(z - \mathbf{W}_{22})^{-1} - z(z - \mathbf{W}_{22})^{-1} = \mathbf{I} - z(z - \mathbf{W}_{22})^{-1}. \tag{22}$$

Taking the trace of this expression gives $N - Nzg_2(z)$. Putting everything together, we finally end up with the algebraic equation

$$\frac{1}{g(z)} = z - 1 + q - qzg(z), \tag{23}$$

where we have defined $q = N/T$. Solving this quadratic equation gives

$$g(z) = \frac{z + q - 1 \pm \sqrt{(z + q - 1)^2 - 4qz}}{2qz} = \frac{z + q - 1 \pm \sqrt{(z - \lambda_+)(z - \lambda_-)}}{2qz}, \tag{24}$$

where we define $\lambda_\pm = (1 \pm \sqrt{q})^2$. It turns out that the correct branch to take of $g(z)$ is the one with the negative radical. Using this in conjunction with Equation 10, one finds that

$$\rho(x) = \frac{\sqrt{(\lambda_+ - x)(x - \lambda_-)}}{2\pi qx} \quad \text{for } \lambda_- < x < \lambda_+. \tag{25}$$

There is an additional subtlety for $N > T$ ($q > 1$). Here our intuition tells us that there should be $N - T = \frac{q-1}{q}N$ trivial zeros of $\mathbf{W}$, since it will not be full rank. Indeed, one can confirm that in this case there is a δ function in the spectral density at $x = 0$. So the full spectral density is

$$\rho(x) = \frac{\sqrt{(\lambda_+ - x)(x - \lambda_-)}}{2\pi q x} + \frac{q-1}{q}\delta(x)\Theta(q-1). \tag{26}$$

3 Covariance Matrices

Do any of these ensembles have any practical relevance? In fact the Wishart ensemble is a very useful tool in signal processing, as it represents a null model for an empirically calculated covariance matrix of a multidimensional time series which is uncorrelated in each of its dimensions. In fact doing principal component analysis on a data matrix $\mathbf{H}$ is equivalent to finding the eigenvalues of the matrix $\mathbf{H}\mathbf{H}^\intercal$, and under the null model that each entry of $\mathbf{H}$ is i.i.d and gaussian, the results of this diagonalization will exactly yield the Marchenko-Pastur distribution. One important question is what happens if the underlying $N \times T$ matrix is no no longer uncorrelated along each of its N rows. This would occur naturally in a situation where different variables are correlated: for example if one measures humidity and temperature separately at a number of different points in time. In this case there is some underlying covariance matrix Σ fore each of the multivariate columns of $\mathbf{H}$. What is the spectral density of $\mathbf{H}\mathbf{H}^\intercal$? In order to answer this question it is necessary to develop some tools of *free probability*, which we will do in subsequent notes.